A Marked Point Process for Modeling Lidar Waveforms

Clément Mallet, Florent Lafarge, Michel J. Roux, Uwe Soergel, Frédéric Bretar, Christian Heipke

A Marked Point Process for Modeling Lidar Waveforms

Clément Mallet, Florent Lafarge, Michel Roux, Uwe Soergel, Frédéric Bretar and Christian Heipke

***Abstract*—Lidar waveforms are 1D signals representing a train of echoes caused by reflections at different targets. Modeling these echoes with the appropriate parametric function is useful to retrieve information about the physical characteristics of the targets. This paper presents a new probabilistic model based on a marked point process which reconstructs the echoes from recorded discrete waveforms as a sequence of parametric curves. Such an approach allows to fit each mode of a waveform with the most suitable function and to deal with both, symmetric and asymmetric, echoes. The model takes into account a data term, which measures the coherence between the models and the waveforms, and a regularization term, which introduces prior knowledge on the reconstructed signal. The exploration of the associated configuration space is performed by a Reversible Jump Markov Chain Monte Carlo sampler coupled with simulated annealing. Experiments with different kinds of lidar signals, especially from urban scenes, show the high potential of the proposed approach. To further demonstrate the advantages of the suggested method, actual laser scans are classified and the results are reported.**

***Index Terms*—Object-based stochastic model, Source modeling, Lidar, Marked point process, Monte Carlo Sampling.**

I. Introduction

A. Lidar remote sensing of topographic surfaces

Airborne laser scanning or lidar (Light Detection And Ranging) is an active remote sensing technique providing direct range measurements between the laser scanner device and the Earth surface. Such distance measurements are mapped into 3D point clouds through a direct georeferencing process involving GPS and inertial measurements [1]. It enables fast, reliable, accurate, but irregular mapping of terrestrial landscapes from geospatial platforms (from satellites to aircrafts). The accuracy of the measurement is high (typically < 0.1 m and < 0.4 m in altimetry and planimetry, respectively).

In remote sensing, laser ranging devices actively emit pulses of short duration (typically a few nanoseconds) in the infrared domain (wavelength between 1 and 1.5 μm) of the electromagnetic spectrum. The distance is derived from the measured round-trip time of the signal between sensor and target. By forward motion of the sensor carrier and an additional scanning mechanism in across-track direction strips of 150 m to 600 m swath width are covered, depending on type of device and carrier altitude.

Due to diffraction, the laser beam inevitably fans out; a typical value for the beam divergence lies between 0.4 and 0.8 mrad. Therefore, a single emitted pulse may cause several echoes from objects located at different positions inside the conical 3D volume traversed by the pulse. This is particularly interesting in forested areas since lidar systems can measure simultaneously both the canopy height and the terrain elevation underneath. Topographic lidar is now fully operational for many specific applications such as metrology, forest parameter estimation, target detection, and power-line, coastal, and open-cast mapping at large scales. 3D point clouds are known to be complementary data to traditional satellite or aerial images as well as hyperspectral data for many issues such as city modeling and building reconstruction [2], and classification of urban or forested areas [3], [4].

The new technology of full-waveform (FW) lidar systems has emerged in the last fifteen years and has become popular the last five years [5]. It permits to record the received signal for each transmitted laser pulse, the result is called a *waveform*. Since the waveform is digitized at constant rate and recorded by the lidar system, FW data is thus a set of equally-spaced discrete samples of the amplitude of the echo signal. Such sample sequence represents the progress of the laser pulse as it interacts with the reflecting surfaces. Hence, FW lidar data yield more than a basic geometric representation of the Earth topography. Instead of clouds of individual 3D points, lidar devices provide connected 1D profiles of the 3D scene, which allows gaining further insight into the structure of the scene. Indeed, each signal consists of series of temporal modes, where each of them corresponds to the reflection from a unique object or a superposition of the signal of several elements (see Figures 1, 2b and 2c).

Since laser scanners with waveform digitizers are becoming increasingly available, many studies have already been carried out to perform advanced signal processing and analysis [5]. The advantage of off-line waveform processing is twofold: by designing his own signal fitting algorithm, traditionally by fitting each echo with a Gaussian curve [6], [7], an end-user can:

(i) Maximize the detection rate of relevant peaks within the waveforms. More points can be extracted in a more reliable and accurate way. Therefore, maximum locations are better determined, and close objects better

C. Mallet, F. Bretar are with the MATIS laboratory, Institut Géographique National, Saint-Mandé, FRANCE – email:firstname.lastname@ign.fr

F. Lafarge is with the ARIANA research group, INRIA, Sophia-Antipolis, FRANCE – email:Florent.Lafarge@inria.fr

M. Roux is with the TSI Department, Telecom ParisTech, Paris, FRANCE – e-mail:michel.roux@telecom.paristech.fr.

U. Soergel, C. Heipke are with the Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover, Hannover, GERMANY – e-mail:lastname@ipi.uni-hannover.de

discriminated [8]. To consistently geolocate the desired reflecting surface, we need to be able to precisely identify the corresponding reflection within the waveform. Such decomposition of the waveforms allows to find the 3D location of the targets.

(ii) Decompose the waveforms by modeling each echo with a suitable parametric function. The echo shape can be retrieved, providing relevant features for subsequent segmentation and classification purposes. Waveform processing capabilities can therefore be extended by enhancing information extraction from the raw signals.

Lidar signal reconstruction is a topic of major interest and a key point for efficient target discrimination. A possible technique is to select for each echo the optimal parametric model taken from a predefined dictionary of modeling functions. This is not a straightforward task, however, and today no automatic techniques for its solution exist. The reason is that the shape of the waveform may vary considerably, and the number of modes is unknown. Their shape can be similar (single-mode) to that of the outgoing pulse, or be complex and multimodal with each mode representing a reflection from an apparently-distinct surface within the laser footprint. Simple waveforms are typical for bare-ground regions and complex waveforms for vegetated areas. Figure 1 enhances the difference between a traditional 3D point cloud and lidar waveform data over a vegetated area, whereas Figure 2 shows some examples of lidar waveforms in various contexts.

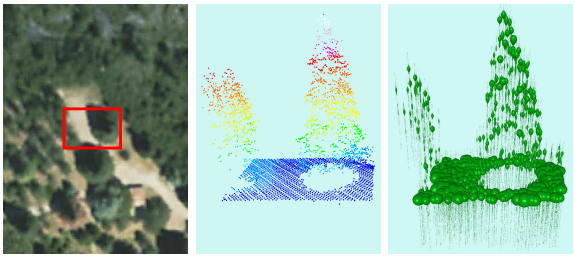


Fig. 1. **Left**: an orthorectified aerial image of a region of interest (ROI – red rectangle) over a vegetated area ©IGN. **Middle**: 3D point cloud of the ROI colored with the altitude (dark blue for lowest altitudes to white ones for highest altitudes). **Right**: Waveforms of the ROI. Each recorded sample of the backscattered signal is represented as a sphere whose radius is proportional to the backscattered energy. The data have been displayed using FullAnalyze [49].

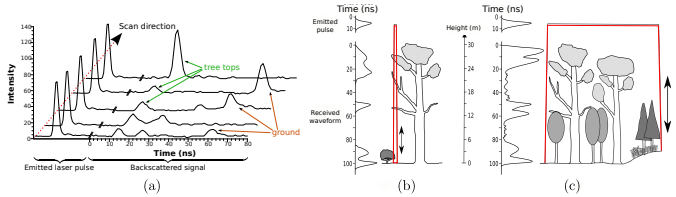


Fig. 2. Some examples of lidar waveforms. **(a)** Successive waveforms plotted in the laser beam direction plane. **(b)** Emitted and received signals in a forested area with a small-footprint lidar (laser beam size at the ground < 1 m). With a small-sized footprint, all targets strongly contribute to the waveform shape, but the laser beam has a high probability to miss the ground. **(c)** Emitted and received signals in a forested area with a large-footprint lidar (size > 5 m). When considering large footprints, the last pulse is bound to be the ground, but each echo is the integration of several targets of identical range at different locations and with different properties.

In this paper, our aim is to model specifically each mode of a lidar waveform by an analytical parametric function.

B. Waveform decomposition as a parameter estimation problem of a Finite Mixture Model

Waveform processing consists in decomposing the waveform into a sum of components or echoes, in order to characterize the different individual targets along the path of the laser beam and model them. On the one hand, methods based on wavelets [9], neural networks [10], splines [11], kernel-based density estimation techniques involving for instance Parzen windows [12] or Support Vector Machines [13] are known to fit 1D signals with large flexibility and efficiency. On the other hand, they do not model each mode of the waveform with the best-fit analytical function of a given a set of parametric curves. Such approach offers two advantages: firstly, the choice of the curve provides insight into the type of interaction involved for modes that result from signal mixture; and secondly, the curve parameters provide additional features for land cover classification.

The problem of finding the best-fit function can instead be addressed by adopting a finite mixture model (FMM) [14] which fulfills our requirements. Mixture models allow us to describe and estimate complex multimodal data by considering them as being sampled from different subpopulations. Indeed, we can postulate the lidar signal to be a linear combination of parametric components, each one corresponding to a specific target. However, the state-of-the-art waveform reconstruction using finite mixture models assumes the mixture component density functions to have a classical parametric form (*i.e.*, Gaussian, uniform, etc.). It should also be noted that many different mixture solutions may explain the same data, and thus, for an interpretability of the mixture, each component should correspond to exactly one mode of the waveform. Historically, estimates of the parameters of the class probability densities in mixture densities have been retrieved via the Expectation-Maximization (EM) algorithm [15], which has found wide application in image and video segmentation. The maximum-likelihood-based method either requires knowledge of the number of components or must be coupled with model selection; many authors have proposed improvements and extensions to this algorithm [16]. Alternatives to EM exist such as Bayesian methods, Kalman filtering, the minimum-distance algorithm, optimization techniques (using, for instance, the gradient descent or the Levenberg-Marquardt algorithm [6]), or the method-of-moments [17]. As opposed to most previous works on FMMs, the model order and the most suitable modeling function for each echo are unknown in our case. Unfortunately, when dealing with parametric functions yielding more complicated analytical expressions, the classical statistical estimation methods fail because their moments do not exist. New approaches have been developed in the Synthetic Aperture Radar (SAR) community to deal with this problem combining the method of log-cumulants and the Mellin transform [18], [19].

C. Motivation

A large body of literature has shown that many remote sensing signals exhibit a more asymmetric nature with heavier tails compared to normal distributions. Also in lidar remote sensing, the Gaussian assumption does not always hold and approximating the waveforms by a sum of Gaussians may be inadequate, depending on the application and the landscape. An emitted laser pulse that interacts with complex natural or man-made objects may cause a multi-echo backscatter sequence of considerable temporal extent. The received power as a function of time can be expressed as follows [7]:

$$P_r(t) = \sum_{i=1}^{C} k_i \; S(t) * \sigma_i(t)\,, \qquad (1)$$

where k_i is a value varying with range between sensor and target, $S(t)$ is the system waveform of the laser scanner and $\sigma_i(t)$ the apparent cross-section of the i^{th} target. S and σ_i are usually described by Gaussian functions, but this is not always correct, and waveforms can be composed of modes with non-similar shapes. To remove both, the broadening and the asymmetric effects caused by a varying S on the received waveforms, a deconvolution step is usually carried out, using for instance matched filtering, Wiener filtering [20], or B-splines. Indeed, target cross-sections are physical parameters which are independent of the emitted pulses. However, such corrections were not introduced in our approach before the modeling step, since asymmetric peaks are also reported after deconvolution. Figure 3 shows waveforms with complex shapes that are different from the Gaussian transmitted pulse. They can be found in the following conditions:

- Two overlapping Gaussian echoes can lead to a single right-skewed pulse (Figure 3a).
- Waveforms acquired with small-footprint sensors (diameter of the laser beam on the ground $\leq$ 1 m) are highly influenced by the local geometry of the intercepted surfaces. They can be positively or negatively skewed by rough surfaces like vegetated areas (trees, hedges) or ploughed fields (Figures 3b and c).
- Waveforms received from large-footprint sensors represent the sum of reflections from all intercepted surfaces illuminated by the conical laser beam (see Figures 2b and 2c). These targets are likely to exhibit a non-symmetric altimetric distribution leading to complex pulse shapes (Figures 3c-d-e).

The traditional approaches dealing with lidar decomposition and modeling [6] are not appropriate for such data. No solution has yet been proposed to transform the well-known parameter estimation problem into an optimal model selection problem for each mixture component where (i) the number of components is unknown and (ii) the parametric models come from a predefined library.

Stochastic methods based on **marked point processes** [21], [22] are very promising for addressing the issue of lidar waveform reconstruction. These models, which allow the sampling of parametric primitives while taking into account complex interactions, have shown very good potential for many applications in remote sensing [23] and especially in image analysis aiming at the extraction of line networks [24], [25], [26], vegetation [27], or 3D urban objects [28], [29]. The sampling of the primitives is performed by Markov Chain Monte Carlo (MCMC) techniques [30] which exhibit very good signal reconstruction properties [31]. Such techniques have been adopted in [32] where a specific model composed of four exponential parametric functions is fitted to lidar intensity histograms of data affected by significant background noise. The model estimate is used for counting and locating the reflected returns from surfaces, as well as retrieving their amplitudes. It thus provides an effective algorithm for 3D ranging, all the more since prior knowledge can also be incorporated into the model. However, this approach is not suitable for our airborne lidar waveform: the parameters of the underlying shape model can vary, but this increases dramatically the dimensions of parameter space and makes the problem much more complex. Thus, the authors of [32] assume all the peaks of the signals to have a similar underlying shape model, an assumption not valid in our case.

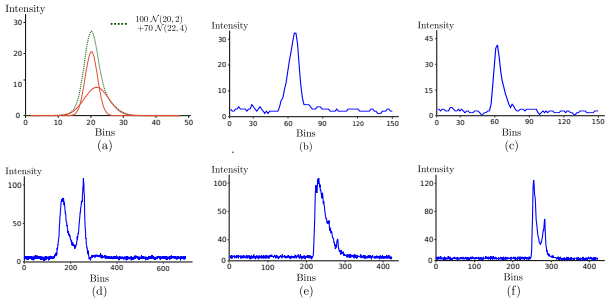


Fig. 3. Types of waveform signals (a) Right-skewed waveform (dark green dashed curve) simulated with two Gaussian pulses (red continuous curves). (b) and (c) Waveforms resulting from the small-footprint laser pulse backscattered from a hedge and a tilted building roof. (d), (e), and (f) Several illustrations of complex asymmetric waveforms acquired on tropical vegetated areas with large-footprint sensors.

This paper presents a method based on a marked point process model that hypothesizes mixtures of various parametric functions representing the reconstructed echos of the airborne lidar waveforms. The optimal configuration of functions is found using a Monte Carlo sampler. Our model presents several interesting characteristics compared to conventional waveform modeling techniques mentioned above:

- *Multiple function types* - The joint sampling of multiple functions types allows to deal with various parametric functions. First, by using a library of shapes, more accurate estimates are performed compared to classical approaches such as the Gaussian mixture model (see [33] and Figure 3). Secondly, by selecting the most suitable function for each peak, which is unknown beforehand, the estimated parameters are more discriminant for a subsequent classification.
- *Lidar physical knowledge integration* - Complex prior information on lidar waveform characteristics can be introduced in the energy of the stochastic model formulation without having problems of convexity or/and continuity restrictions in the formulation of these interactions. This permits to get a more realistic model and to achieve better

results.

- *Efficient exploration of configuration spaces* - A MCMC sampler associated with relevant proposition kernels allows us to avoid exhaustive explorations of large configuration spaces, which can be both continuous and discrete. It is particularly efficient when the number of functions is unknown.

Thus, the Reversible Jump Markov Chain Monte Carlo (RJMCMC) [34] algorithm is attractive because in a multi-object framework it can deal with parameter estimation and model selection jointly in a single paradigm.

This paper extends the work we presented in [35] by improving the model, detailing both the marked point process and the optimization technique, and by presenting new results from various kinds of sensors as well as applications to the classification of urban areas. Section II introduces marked point processes. The proposed model is formulated in Section III. Section IV describes the optimization procedure. Results are shown in Section V including experiments from various kinds of sensor data showing the flexibility of our approach. The application of waveform modeling for image classification in urban areas is also presented. It underlines the good potential of our approach. Finally, conclusions are drawn and perspectives for further work are given in Section VI.

II. Marked point processes

The marked point processes are stochastic tools which have been introduced in signal and image processing by Baddeley and Van Lieshout [21], and extended further in [22], [36], [37]. These models can be considered as an extension of conventional Markov Random Fields [38] such that random variables are associated not with signal values but with parametrical functions describing the signal. An overview of marked point processes is given below.

A. Point processes

Let us consider X, a *point process* living in a continuous bounded set $K = [0, L_{\max}]$ supporting a 1D signal. X is a measurable mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to the set of configurations of points of K:

$$\forall\, \omega \in \Omega,\ x_i \in K,\ X(\omega) = \{x_1, ..., x_{n(\omega)}\}\,, \tag{2}$$

where $n(\omega)$ represents the number of points associated with the event ω. The *homogeneous Poisson process* is the reference point process. Let $\nu(.)$ be a positive measure on K. A Poisson process X with intensity $\nu(.)$ possesses the two following properties:

- For every Borel set $B \in K$, the random variable $N_X(B)$ defining the number of points of X in the Borel set B follows a discrete Poisson distribution with the mean $\nu(B)$, *i.e.*:

$$P(N_X(B) = n) = \frac{\nu(B)^n}{n!}\, \mathrm{e}^{-\nu(B)}.$$

- For every finite sequence of non intersecting Borelian sets $B_1, ..., B_l$, the random variables $N_X(B_1), ..., N_X(B_l)$ are independent.

The Poisson process induces a complete spatial randomness, given by the fact that the positions are uniformly and independently distributed. Its role is analogous to Lebesgue measures on $\mathbb{R}^d$.

B. Density and Gibbs energy

Complex point processes introducing both, consistent measurements with data and interactions between points, can be defined by specifying a density with respect to the distribution of a reference Poisson process. Let us consider an homogeneous Poisson process with intensity measure $\nu(.)$ and let $h(.)$ be a non-negative function on the configuration space $\mathcal{C}$. Then, the measure $\mu(.)$ having a density $h(.)$ with respect to $\nu(.)$ is defined by:

$$\forall B \in \mathcal{B}(\mathcal{C}),\ \mu(B) = \int_B h(x)\nu(dx)\,. \tag{3}$$

A Gibbs energy $U(x)$ can also be used to specify a point process. The density $h(x)$ of a configuration x is then formulated using the Gibbs equation:

$$h(x) = \frac{1}{Z}\, \mathrm{e}^{-U(x)}\,, \tag{4}$$

where Z is a normalizing constant such that $Z = \int_{x\in\mathcal{C}} \mathrm{e}^{-U(x)}$. When defining the Gibbs density of the associated marked point process w.r.t. the Poisson measure, the issue is reduced to an energy minimization problem. Generally, a Monte Carlo Markov Chain sampler coupled with a simulated annealing is used to find the maximum density estimator[1] $\widehat{x} = \arg\max h(.)$. This optimization process is particularly interesting since the density $h(.)$ does not need to be normalized. Thus, the complex computation of the normalizing constant Z is avoided.

C. Marks and object library

In order to model signals in terms of parametric functions, it is possible to extend a point process by adding specific *marks* that associate a parametric function (also called an *object*) to each point[2]. A *marked point process* in $S = K \times M$ is a point process in K where each point is associated with a mark from a bounded set M (see Figure 4).

Usually, the marked point process based models [28], [24], [25], [26], [39] use a single type of object. Some authors [40] have extended the conventional framework in order to sample various kinds of objects extracted from a library. The mark space M associated with this library is then specified as a finite union of mark bounded subsets M_q:

$$M = \bigcup_{q=1}^{N_s} M_q\,, \tag{5}$$

where each subset M_q corresponds to one of the N_s specific object types. This extension of the marked point processes, which is able to deal with objects having different numbers of control parameters, will be used in the following.

[1]This estimator corresponds to the configuration minimizing the Gibbs energy $U(.)$, *i.e.*, $\widehat{x} = \arg\min U(.)$

[2]In many cases, the point corresponds to the mean of the function.

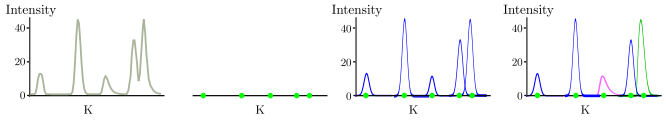


Fig. 4. Illustration of various point processes - **From left to right**: a 1D signal defined on the support K, realizations of a point process on K, a marked point process of Gaussian functions, and a point process specified by a library of various functions.

III. Stochastic model formulation

A. Library of modeling functions

As underlined in Section I-C, the contents of the library is a key point in our work since the function parameters will be used subsequently for classifying the lidar data. Three different distributions are chosen to model the waveforms. Their parameters are defined in continuous domains.

The Gaussian and **Generalized Gaussian** (GG) models have been shown to fit most of the echoes of small-footprint lidar waveforms in urban areas [7]. They allow to model symmetric echoes which form the majority of lidar signals. The GG function can be expressed as follow:

$$f(x \mid I, s, \alpha, \sigma) = I \, \exp\left(\frac{-(x-s)^{\alpha^2}}{2\sigma^2}\right), \qquad (6)$$

where I and σ give the amplitude and the width of the Gaussian model, which are traditionally integrated in lidar classification algorithms. It was shown that they are relevant features for classification in urban areas [41]. A shape parameter α is added to cope with distorted symmetric echoes. It enables to simulate traditional Gaussian shapes when $\alpha = \sqrt{2}$, more peaked curves when $1 \leq \alpha < \sqrt{2}$ ($\alpha = 1$ gives the Laplace function), and flattened shapes when $\alpha > \sqrt{2}$. Shift parameter s was introduced to indicate the position of the maximum of the function.

Nevertheless, the Gaussian assumption does not always hold. Non-unique asymmetric echoes are observed within waveforms corresponding to ground surface or tree canopy (Figure 3). Thus, many waveforms exhibit heavier tails and require a more flexible parametric characterization. Moreover, the GG model gives the amplitude, width, and shape for symmetric echoes. Amplitude and width are useful for discriminating ground, vegetation, and buildings, but fail to segment different kinds of surfaces such as grass, gravel, and asphalt, even when the pulse shape is available [42]. The laser cross-section gives slightly better discrimination.

Two kinds of functions must therefore be included: functions able to fit asymmetric peaks and those which can cope with both left- and right-skewed curves which therefore deliver other parameters than those provided by the GG model: the **Nakagami** and the **Burr** models have been selected.

The **Nakagami** distribution is a generalization of the χ distribution and can model right-skewed and left-skewed distributions with a skewness/spread parameter ω:

$$f(x|\, I, s, \xi, \omega) = I \, \frac{2\,\xi^{\xi}}{\omega \Gamma(\xi)} \left(\frac{x-s}{\omega}\right)^{2\xi-1} \exp -\xi \left(\frac{x-s}{\omega}\right)^{2}. \qquad (7)$$

When ω increases, the peak becomes narrower and more symmetric. Scale parameter ξ controls the peak width: large ξ leads to narrow peaks of higher amplitude. The Nakagami function is traditionally used to model Synthetic Aperture Radar (SAR) images to estimate their amplitude probability density functions as well as for subsequent classification [18]. A large body of literature has presented and studied probability density functions so as to model the dispersion of the received signals produced by different objects, using either theoretical or heuristic models [43], [19].

Finally, the **Burr** function is especially useful to model asymmetric modes with two shape parameters. It enables to fit right-skewed peaks that the Nakagami model cannot handle. It is a generalization of the Fisk distribution thanks to the parameter c. The scale parameter is a, and b and c are two shapes parameters (b has the same effect as the ω parameter for the Nakagami function). The ratio between peak amplitude and skewness is tuned by c.

$$f(x|\, I, s, c, a, b) = I \, \frac{bc}{a} \left(\frac{x-s}{a}\right)^{-b-1} \left(1 + \left(\frac{x-s}{a}\right)^{-b}\right)^{-c-1}. \qquad (8)$$

On the one hand, we admit that there is no physical entity exclusively attached to these curves. On the other hand, they enable us to handle asymmetric peaks and therefore we expect their application will outperform standard approaches. These distributions are defined in continuous domains. Table I provides some representations of these functions with critical parameter variations.

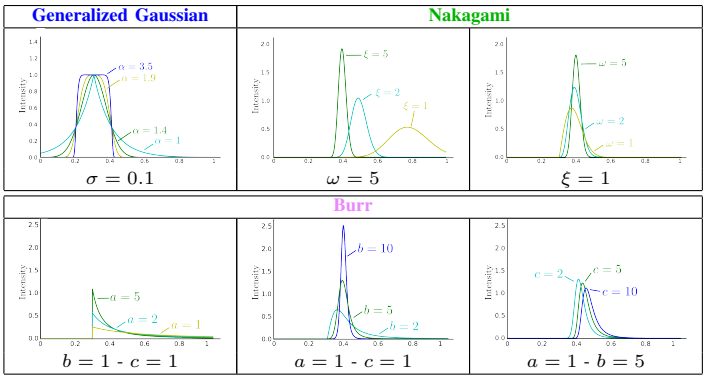


TABLE I
Behavior of the three modeling functions of the library.

B. Energy definition

Let x be a configuration of parametric functions (or objects) x_i extracted from the above library. The energy $U(x)$, measuring the quality of x, is composed of both a data term $U_{\mathrm{d}}(x)$ and a regularization term $U_{\mathrm{p}}(x)$ such that:

$$U(x) = (1-\beta)\, U_{\mathrm{d}}(x) + \beta\, U_{\mathrm{p}}(x), \qquad (9)$$

where $\beta \in \mathbb{R}^+$ tunes the trade-off between the data term and the regularization.

1) Data term: The data energy steers the model to best fit to the lidar waveforms. The likelihood can be obtained by computing a distance between the given signal $\mathcal{S}_{\text{data}}$ and the

estimated one $\mathcal{S}_x$, which depends on the current objects on the configuration x:

$$U_{\mathrm{d}}(x) = \sqrt{\frac{1}{|K|}\int_K (\mathcal{S}_x - \mathcal{S}_{\mathrm{data}})^2}\,. \qquad (10)$$

The term $U_{\mathrm{d}}(x)$ measures the quadratic error between both signals: it allows to be sensitive to high variations *i.e.*, to local strong errors in the signal estimate that correspond to unfitted peaks. The L_2 norm has been chosen for that purpose.

2) Waveform constraints: The term $U_{\mathrm{p}}(x)$ allows the introduction of interactions between objects of x and to favor/penalize some configurations.

$$U_{\mathrm{p}}(x) = U_n(x) + U_e(x) + \sum_{x_i \sim x_j} U_m(x_i, x_j)\,, \qquad (11)$$

where $x_i \sim x_j$ constitutes the set of neighboring objects in the configuration x. This neighborhood relationship $\sim$ is defined as follow:

$$x_i \sim x_j = \{(x_i, x_j) \in x \quad | \quad |\mu_{x_i} - \mu_{x_j}| \leq r\}\,. \qquad (12)$$

Parameter μ_{x_i} (resp. μ_{x_j}) represents the mode (*i.e.*, the position of the maximum amplitude of the echo) of the associated function to object x_i (resp. x_j) and r is constrained by the lidar sensor range resolution (*i.e.*, the minimum distance between two objects along the laser line of sight that can be differentiated) as well as the complexity of the reconstruction we aim to achieve.

For aerial lidar waveforms the prior knowledge is set up by physical limitations in the backscatter of lidar pulses. These limitations are modeled by three terms U_n (echo number limitation), U_e (backscatter laser energy limitation), and U_m (reconstruction complexity) that are described below.

(i) Echo number limitation

The two first echoes of a waveform contain in general about 90% of the total reflected signal power. Consequently, even for complex targets like forested areas, a waveform empirically reaches a maximum of seven echoes and it is quite rare to find more than four echoes. In urban areas, most of the targets are rigid, opaque structures like buildings and streets. Thus, more than two echoes are usually only found in open forests. We therefore aim to favor configurations with a limited number of objects with an energy given by:

$$U_n(x) = -\log \mathrm{P}_{\mathrm{card}(x)} \quad \text{with} \quad \sum_{n=0}^{\infty} \mathrm{P}_n = 1\,, \qquad (13)$$

where P_n is the probability for the waveform to have n echoes. The probabilities were empirically determined by a coarse mode estimate on an urban test area (41M waveforms over 20 km^2). Here, we have: $P_1 = 0.6$, $P_2 = 0.27$, $P_3 = 0.1$ and $P_{4 \leqslant n \leqslant 7} = 0.01$. For $n > 7$, $U_n(x)$ is set to a very high positive value, which bans such configurations in practice.

(ii) Backscatter energy limitation

We take advantage of the law of conservation of energy and define an upper bound for the backscatter energy. This upper bound depends on the emitted laser power and the target reflectance and scattering properties. This reference power E_{ref} can be set empirically to $\sqrt{2\pi} A_{\max} \sigma_{\max}$, which is the energy of a Gaussian pulse of amplitude $A_{\max}$ and width $\sigma_{\max}$. $A_{\max}$ and $\sigma_{\max}$ are upper bounds for the amplitude and the width of echoes within the waveforms over the area of interest. Waveforms with larger pulse energy are penalized as follows:

$$U_e(x) = \pi_e \, 1_{\{E(x) > E_{\mathrm{ref}}\}} (E(x) - E_{\mathrm{ref}})^2\,, \qquad (14)$$

where $1_{\{.\}}$ is the characteristic function, $E(x) = \int_K \mathcal{S}_x$ is the pulse energy of $\mathcal{S}_x$, compared to a reference power E_{ref} (see Figure 5).

(iii) Reconstruction complexity

Our aim is twofold:

- to penalize objects spatially closer along the line of sight than the sensor range resolution;
- to favor configurations with a small number of objects, following the Minimum Description Length principle.

Such energy is given by:

$$U_m(x_i, x_j) = \pi_m \exp\left(\frac{r^2 - |\mu_{x_i} - \mu_{x_j}\,.\,|^2}{\sigma^2}\right) \qquad (15)$$

This means that a mode of a waveform may be either reconstructed by a single peak or by a sequence of peaks whose accepted minimum distance is governed by parameter r (see Figure 5). The lower bound of r is given by range resolution $\tau \times c/2$ (where τ is the laser pulse duration, and c le speed of light), while the upper bound of r is thus model based and may be chosen depending on the scene. For example, if we know that the data were acquired in a forested area in the leaf-off period and the trees have preferably few, but strong branches, we would chose a large r.

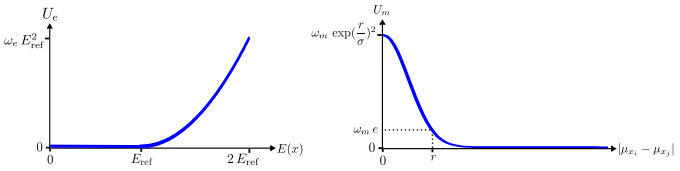


Fig. 5. **Left**: Backscatter energy limitation term plotted against the energy of the current configuration $E(x)$. **Right**: Reconstruction complexity term plotted against the absolute distance between two neighboring objects of the current configuration.

3) Parameter settings: Physical and weight parameters can be distinguished in the energy. Physical parameters are r and σ. Small-footprint airborne topographic sensor specifications [5] and our knowledge on acquired waveforms lead to $r = 0.75$ m, and we set σ to 0.01. Thus, $R_3(x_i, x_j) \to +\infty$ when $\mu_{x_i} \to \mu_{x_j}$. Data and regularization terms are weighted with respect to each other using a factor β (see Equation 9) set to 0.5. The two prior weights π_e and π_m are tuned by "trial-and-error" tests.

IV. Optimization by Monte Carlo sampler

We aim to find the configuration of objects which minimizes the non convex energy U in a variable dimension space since

the number of objects is unknown and function types are defined by different numbers of parameters. Such a space can be efficiently explored using a Monte Carlo sampler coupled with a simulated annealing that we detail below.

A. MCMC sampler

Since it is required to sample from parameter spaces of varying dimensions, the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm [34] is well adapted to our problem. This technique is a general extension of the formalism introduced in [30] for variable dimension models. [34] proposes a selection of models in cases of a mixture of k Gaussian, since k is not known. Several papers have shown the efficiency of the RJMCMC sampler for the problem of multiple parametric object recognition [44], [25], [45] in image processing and computer vision.

The RJMCMC sampler consists in simulating a discrete Markov Chain (X_t), t $\in \mathbb{N}$ on the configuration space, having an invariant measure specified by the energy U. This sampler performs "jumps" between spaces of different dimensions respecting the reversibility assumption of the Markov chain. One of the advantage of this iterative algorithm is that it does not depend on the initial state. The jumps are realized according to various families of moves m called proposition kernels and denoted by Q_m. The jump process performs a move from an object configuration x to y according a probability $Q_m(x \rightarrow y)$. Then, the move is accepted with the following probability:

$$\min\left(1,\ \frac{Q_m(y \rightarrow x)}{Q_m(x \rightarrow y)}\ \exp -(U(y) - U(x))\right). \qquad (16)$$

Two families of moves are used in order to perform jumps between the subspaces. Another type of move is more specifically dedicated to the exploration of such subspaces.

- **Birth-and-death kernel** Q_{BD}: an object is added or removed from the current configuration x, following a Poisson distribution. These transformations corresponding to jumps into the spaces of higher (birth) and lower (death) dimension are theoretically sufficient to visit the whole configuration space. However, other kernels, more adapted to our problem, can be specified. The aim is to speed up the process convergence by proposing relevant configurations more frequently. Therefore, two other kernels have been introduced.
- **Perturbation kernel** Q_P: the parameters of an object belonging to the current configuration x are modified according to uniform distributions.
- **Switching kernel** Q_S: the type of an object belonging to x is replaced by another type of the library. Contrary to the previous kernel, this move does not change the number of objects in the configuration. However, the number of parameters can be different (*e.g.*, four parameters for the Nakagami model are substituted by five parameters for the Burr one). This kernel creates bijections between the different types of objects [34].

If an object is added, its type and its associated parameters are randomly chosen. Because no assumption can be made which move is more relevant at the current state, we choose equiprobability of the kernels in order to not favor one with respect to another. The computation of these kernels is detailed in Appendix B.

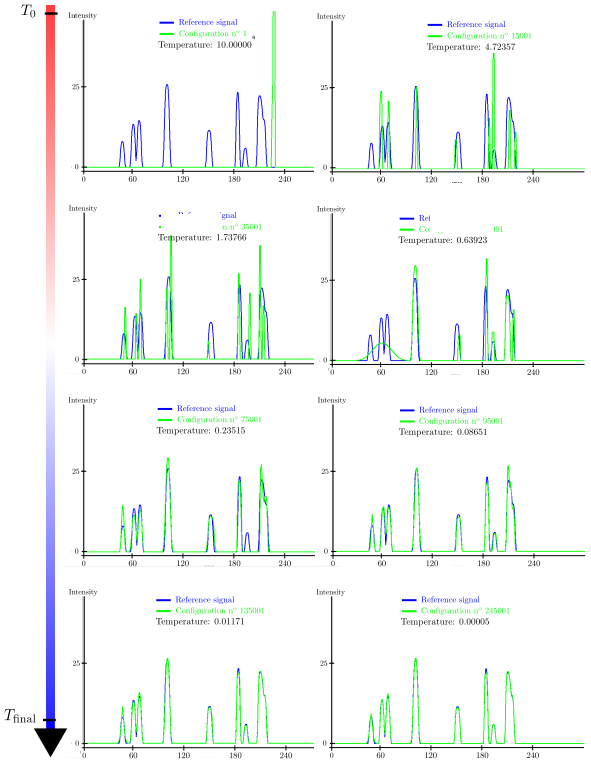


Fig. 6. Optimization process: evolution of the object configuration from the initial temperature T_0 to the final one T_{final} (left to right, and top to down).

B. Relaxation

Simulated annealing is used to ensure the convergence process. A relaxation parameter T_t, defined by a sequence of temperatures decreasing to zero when $t \rightarrow \infty$, is introduced in the RJMCMC sampler (*i.e.*, $U(.)$ is substituted by $\frac{U(.)}{T_t}$). Simulated annealing allows to theoretically ensure the convergence to the global optimum for all initial configurations x_0 using a logarithmic temperature decrease. In practice, we prefer to use a geometrical cooling scheme which is faster and gives an approximate solution close to the optimal one:

$$T_t = T_0\ \alpha^t\,, \qquad (17)$$

where α and T_0 are the decrease coefficient and the initial temperature, respectively. We prefer to use a constant decrease coefficient. In our experiments, α is set to 0.99995. The initial temperature T_0 is estimated according to [46]. During the temperature decrease, the process explores the configurations of interest and becomes more and more selective. It corresponds to local adjustments of the objects of the configuration (see Figure 6).

V. Experiments

The algorithm has been applied to different kinds of airborne lidar signals. The results have been evaluated quanti-

tatively by computing the normalized cross-correlation coefficient ρ and the relative Kolmogorov-Smirnov distance KS between the raw and the estimated signals. We have:

$$\rho = \frac{\sum_{i=1}^{N} (\mathcal{S}_{\text{data}}(i) - \overline{\mathcal{S}_{\text{data}}}) \cdot (\hat{\mathcal{S}}(i) - \overline{\hat{\mathcal{S}}})}{\sqrt{\sum_{i=1}^{N} ((\mathcal{S}_{\text{data}}(i) - \overline{\mathcal{S}_{\text{data}}})^2} \sqrt{\sum_{i=1}^{N} (\hat{\mathcal{S}}(i) - \overline{\hat{\mathcal{S}}})^2}} \quad \in [\text{-}1,1]\,. \tag{18}$$

$\mathcal{S}_{\text{data}}$ is the reference waveform, and $\hat{\mathcal{S}}$ is our estimated signal, both composed of N bins. $\overline{\mathcal{S}_{\text{data}}}$ and $\overline{\hat{\mathcal{S}}}$ are their respective mean values. If the reconstructed signal perfectly fits with the lidar waveforms bins, ρ=1. The correlation coefficient is rather sensitive to outliers. KS is a normalized L_∞ norm, both used to detect missing echoes and local shifts between signals. It is defined as follows:

$$\text{KS}(\mathcal{S}_{\text{data}}, \hat{\mathcal{S}}) = \frac{\sup_K |\mathcal{S}_{\text{data}} - \hat{\mathcal{S}}|}{\max_K \mathcal{S}_{\text{data}}} \quad \in [0,1]\,. \tag{19}$$

The L_∞ norm has been normalized to allow comparisons between waveform fitting results from different sensors. KS=0 means that every lidar bin perfectly matches with the reconstructed signal, whereas KS=1 means that the main echo has been missed. Setting a satisfactory KS upper bound thus mainly depends on the noise level of the lidar waveforms.

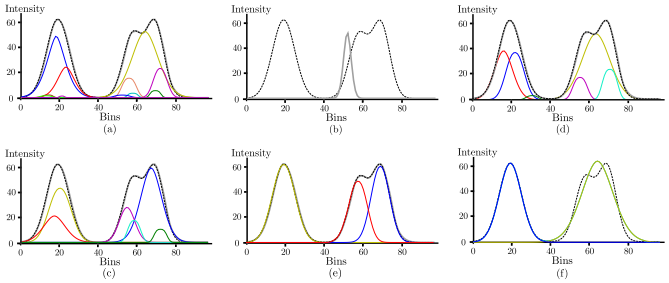


Fig. 7. Various signal reconstructions with variations of the model energy. The dotted black line and the continuous grey one are respectively the raw and the estimated signals. The other colors correspond to the individual echoes that compose the estimated waveform. (a) No regularization term. (b) No data term. Both data and regularization terms are now considered (first, U_n and U_e). The reconstruction complexity term U_m is not used in (c). U_m is introduced and then modified by increasing the parameter r: r is set to 0.3, 0.75, and 3 m respectively for (d), (e), and (f).

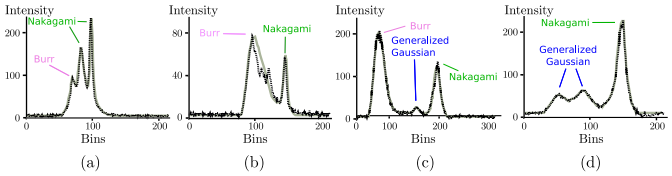


Fig. 8. Examples of fitted **(a-b)** LVIS and **(c-d)** SLICER waveforms. The Burr and Nakagami models are preferred for the first and last modes of the waveform, that correspond to the first layer of the three canopy and the ground, respectively. Waveform **(b)** has been fitted setting r to a high value, to prevent small overlapping echoes to be individually detected.

A. Simulated data: relevance of the optimized energy

Various simulations have been carried out to assess the relevance and the effectiveness of the proposed model energy. Figure 7 shows several reconstructions of a simulated signal composed of three pulses with two overlapping peaks with variations on the optimized energy. The simulated signal and the estimated one are represented by the dotted black line and the continuous grey line, respectively.

First, the data term has been considered only (*i.e.*, we only minimize the difference according to Equation 10). It can be noticed in Figure 7a that the signal is correctly estimated but the configuration is composed of a high number of echoes (eleven). Thus, the result is not realistic since not all echoes does represent a specific target. Then, only the regularization term is considered ($U_n + U_e + U_m$). Figure 7b shows that the proposed regularization energy constraint is useful since it provides a realistic lidar waveform: one echo with a bounded energy. This is due to both, the echo number limitation and the backscatter energy limitation terms. Finally, Figures 7c to 7f show the influence of the reconstruction complexity term U_m. It is first discarded on Figure 7c: the signal is perfectly reconstructed, but with the maximum number of echoes allowed by U_n. It does not correspond to reality since the echoes are too closely located to each other. Then, U_m is introduced and r is respectively set to 0.3, 0.75, and 3 m in Figures 7d, 7e, and 7f. The greater r, the lower the number of peaks. It can be noticed that a reasonable value of r allows the reconstruction of the signal with the appropriate number of echoes (Figure 7e), whereas larger values lead to erroneous detections (Figure 7f).

Additional experiments have been carried out to assess whether the parameters of the modeling functions are correctly estimated and whether the correct model is selected (see Appendix A for details).

B. Airborne medium and large-footprint topographic waveforms

Waveforms from LVIS (Laser Vegetation Imaging Sensor) and SLICER (Scanning Lidar Imager of Canopies by Echo Recovery) NASA sensors have been decomposed and modeled with our approach (Figure 8). The sensor goals and specifications are described in [5]. LVIS waveforms have been acquired in March 1998 over a 800 km^2 area of Costa Rica using 25 m-diameter footprints[3] [47]. Both fine and coarse fitting strategies have been tested. The fine strategy consists in selecting r so that each mode of the waveform will be fitted by a function ($r = 3$ m). It leads to almost perfect signal approximation, but conclusions are difficult to draw since the function selected for a given peak depends on the functions of the neighboring echoes (Figure 8a). With the coarse solution, r is set to higher value (9 m) and σ to 0.001. Thus reduces the complexity of the reconstruction and therefore prevents overlapping or close echoes from being individually fitted. A unique global peak is selected instead (Figure 8b), providing a general trend for the first part of the signal (in practise, the first tree canopy layer). SLICER elevation profiles come from in the BOREAS Northern Study Area in Canada[4], and have been acquired in

[3]Data set available at https://lvis.gsfc.nasa.gov/index.php

[4]Data set available at http://core2.gsfc.nasa.gov/research/laser/slicer/browser.html

Sensor (number of waveforms)	ρ	KS	GG	Nak.	Burr
SLICER (76417)	0.949	0.11	6.5	51.2	41.8
LVIS (4001)	0.968	0.14	5.1	57.0	37.9

TABLE II
MEDIUM AND LARGE-FOOTPRINT WAVEFORM FITTING AND MODELING STATISTICS. THE FINE SOLUTION HAS BEEN ADOPTED FOR THE SIGNAL DECOMPOSITION. THE TWO FIRST COLUMNS (ρ – KS) PROVIDE QUALITY MEASURES. THE THREE LAST COLUMNS INDICATE THE PERCENTAGE OF ECHOES THAT HAVE BEEN FITTED WITH EACH OF THE THREE MODELING FUNCTIONS.

July 1996 [48]. Table II shows that signals from both sensors are correctly decomposed, without significant errors. However, the KS distance values show that some small peaks are not retrieved. Indeed, LVIS and SLICER elevation profiles are very complex since the sensor laser beam integrates many distinct objects. Thus, even with the fine strategy, several close narrow pulses (as displayed on Figure 8b) cannot be all detected.
With medium and large-footprint waveforms, the Generalized Gaussian model is no longer selected by the algorithm. The two functions allowing to simulate asymmetric peaks are preferred. The main noticeable results (see also Table II) are that:

- the GG function is sparsely chosen, mainly for peaks with a small amplitude. Thus, the lidar echo Gaussian assumption is no longer valid. This fact underlines the relevance of our approach for modeling lidar waveforms with a library of functions.
- the Nakagami model is preferred to the Burr function, since its parameters allow for a higher flexibility. It is mainly selected for the last echo, which correspond to the ground and low above-ground objects, and is usually left-skewed.
- the Burr function is relevant for echoes that correspond to pulses backscattered from the tree canopy (first layer of the vegetation).

C. Small-footprint waveforms in urban areas

Waveforms acquired from small-footprint airborne lidar systems (Riegl LMS-Q560 and Optech 3100EA, see [5] for their specifications) over various kinds of urban landscapes have been fitted using the stochastic approach. Figures 9 and 10, and Table III show results both on urban and natural terrain. First, it can be noticed in Figure 9 that the algorithm performs well on complex waveforms. The correct number of echoes is found as well as the correct shape of the waveform: single and multiple overlapping echoes are retrieved, even in vegetated areas where the noise level is significant w.r.t. the echo amplitudes (Figures 9a and b). Moreover, for opaque solid targets like building roofs and ground, slightly asymmetric echoes are retrieved, and correctly adjusted: the Burr model allows to retrieve them, especially when dealing with the second echo of two overlapping ones (Figures 9c and d).
The fitting accuracy is higher than for medium and large-footprint waveforms. However, the latter ones are much more complex.

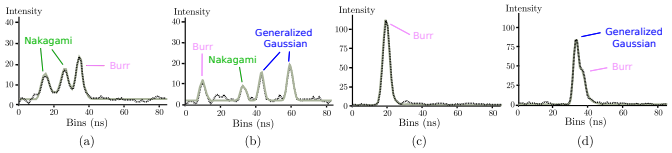


Fig. 9. Decomposed and modeled waveforms on **(a-b)** trees, **(c)** a building roof, and **(d)** a hedge (Riegl LMS-Q560 sensor).

More than 123,000 waveforms acquired with the Optech 3100EA sensor over the city of Amiens, France, have been analyzed. The aim was to assess the reliability of the method in heterogeneous landscapes and to show its local stability in homogeneous areas. Six regions of interest have been selected: three simple buildings with different slopes and materials (*Building #1*); a complex area with high and low buildings with grass and trees (*Building #2*); a Gothic cathedral (*Cathedral*); a flat harvested field (*Field*); a slightly sloped grass surface (*Grass*); and a mixed set of buildings with a street, pavement, and trees (*Street*). Furthermore, the echoes detected by the lidar system during the acquisition survey are provided (*hardware* echoes). To assess the relevance of waveform processing and to solve a multiple mixture problem, the waveforms have been fitted with our proposed approach using the library of functions as well as based exclusively on the Gaussian model. All the results are included in Table III. The main conclusions are:

- For areas including targets generating multiples echoes (trees, building edges) more echoes are found than the traditional 3D point cloud provides to the end-user. These areas correspond approximatively to 5-10% for urban scenes as stated in several papers in the literature [49].
- Whatever the ROI, the fitting accuracy is high (ρ >0.99 and KS<0.1) with our approach. One can notice that using a library of shapes slightly improved the fitting accuracy compared to only the Gaussian model (since we have gained a higher flexibility in the fitting process with new models featuring more and distinct parameters). There are indeed asymmetric peaks, but in a relative low proportion.
- For flat areas, which coincide with low incidence angles, the echoes are symmetric and the Generalized Gaussian function is selected (*Field* and *Grass* areas, see Figure 10). However, in some cases also the Burr and Nakagami functions have been selected because for some parameter set-ups they are very similar to Gaussian distributions. This is a limitation of the current version of our approach which will be targeted in future work.
- In vegetated areas (trees), the algorithm does not preferably select a particular model. The usefulness of asymmetric modeling functions is therefore difficult to draw for fitting echoes of small-footprint waveforms in forested areas, and the Gaussian function should be sufficient.
- For building regions, both symmetric and skewed peaks are retrieved. Asymmetric echoes can be found on building roofs and where surface discontinuities exist. Such behavior is frequently observed for the Cathedral scene depicted in Figure 10. When the target geometry becomes

Area		Building #1	Building #2	Cathedral	Field	Grass	Street
# waveforms		9943	38565	43563	10035	9790	11770
# echoes hardware		10555	40785	49161	10035	9790	12428
Library	# echoes	11054	43385	50638	10035	9790	13033
	GG	81.2	60.3	62.5	91.2	99.5	61.7
	Nak.	12.6	35.3	27.4	4.4	0.35	31.4
	Burr	6.2	4.4	10.1	4.2	0.15	6.9
	ρ	0.9947	0.9948	0.9948	0.997	0.999	0.994
	KS	0.098	0.0977	0.095	0.038	0.025	0.102
Gauss.	ρ	0.991	0.987	0.9824	0.992	0.995	0.981
	KS	0.109	0.125	0.113	0.087	0.057	0.134
Additional points		+4.7%	+6.4%	+3%	0%	+0.5%	+4.8%

TABLE III
FITTING RESULTS ON SIX URBAN REGIONS OF INTEREST (OPTECH 3100EA SENSOR). OUR APPROACH HAS BEEN TESTED USING BOTH THE FULL LIBRARY OF MODELS AND THE SINGLE GAUSSIAN FUNCTION. QUALITY MEASURES (ρ – KS) ARE PROVIDED FOR BOTH. THE PERCENTAGES OF ECHOES THAT HAVE BEEN FITTED BY EACH OF THE THREE MODELING FUNCTIONS ARE INDICATED, AS WELL AS THE PERCENTAGE OF ECHOES ADDITIONALLY RETRIEVED, COMPARED TO THE UNKNOWN HARDWARE DETECTION METHOD.

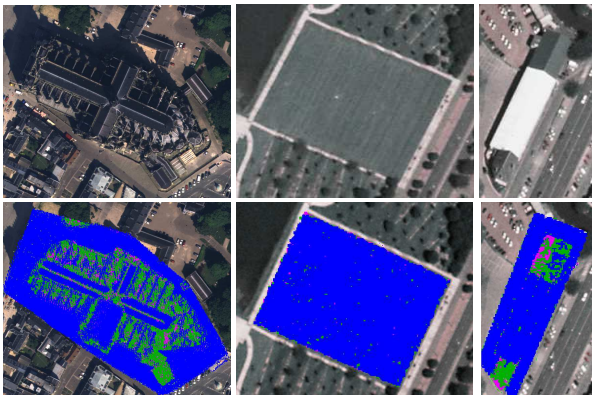


Fig. 10. **From top to down**: orthoimages of the ROIs ©IGN, and 3D point clouds interpolated in 2D and colored with the selected function: Generalized Gaussian–Nakagami–Burr. **Left**: Cathedral – **Middle**: Grass – **Right**: Building #1.

complex, the Nakagami and Burr functions are preferred.

- The reflectance of the targets also has an influence on the fitting algorithm: for high reflectance objects, the backscattered pulse has a significant amplitude and becomes narrower. In such cases, the Gaussian model is selected, as displayed in Figure 10 for the *Building #1* area.

D. Application to lidar data classification

1) Motivation and strategy: A potential application is data classification using the modeling features. The aim is to assess whether such features are relevant for accurate urban land cover classification. These features can be fed into a classification algorithm using for instance Support Vector Machines (SVM). SVM have evolved as a standard tool for a broad range of classification tasks [50], [51].

Here, our goal is to carry out simple classification without selecting the most relevant features. Three classes have been chosen to characterize urban areas: **buildings**, **vegetation**, and **ground**. Moreover, with such coarse classes, a 2D-based classification is preferred. 3D lidar points are thus projected into a 2D image geometry (0.75 m resolution). Images are obtained for each feature by computing, for each pixel, the mean corresponding value of the lidar points included in a 3×3 neighborhood. Such interpolation process has been proven to be efficient for classifying lidar data with few errors on class boundaries [41]. The SVM algorithm requires a feature vector for each pixel to be classified. Our feature vector f_v has **eight** components. Four of those are **spatial features**, which are computed using a volumetric approach within a local neighborhood $\mathcal{V}_P$ for each lidar point P. The local neighborhood includes all the lidar points within a sphere of a fixed radius (set to 2 m), centered at P. Four other features are extracted from the waveform processing step (**shape features**). Finally, $f_v = [\Delta z, \sigma_z, D_\Pi, S_\lambda, A, w, s, \mathcal{M}]$.

- Δz: difference between the echo altitude and the lowest altitude in a neighborhood of 20 m;
- σ_z: the variance of the altitude of the points found in $\mathcal{V}_P$;
- D_Π: distance from the current point P to the locally estimated plane Π_P. Such plane is estimated using a robust M-estimator with $L_{1.2}$ norm;
- S_λ: the sphericity is equal to λ_3/λ_1. λ_1 and λ_3 are the highest and lowest eigenvalues, respectively extracted from the covariance matrix computed in $\mathcal{V}_P$;
- the peak amplitude A, width w, skewness s, and the type selected by the marked point process $\mathcal{M}$.

2) Results and comments: The six ROIs have been classified. To assess the relevance of the features extracted with the modeling step, the classification has first been carried out using only the echo shape features: $f_{v_1} = \{A, w, s, \mathcal{M}\}$. Then, the four spatial features were selected: $f_{v_2} = \{\Delta z, \sigma_z, D_\Pi, S_\lambda\}$. Finally, the four shape features are successively introduced into f_{v_2}. The overall accuracy (OA) is used as a quality criterion for comparing the results and is defined as:

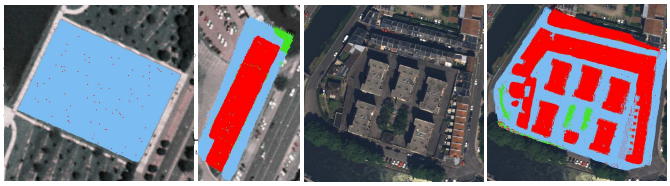
$$\text{OA} = \frac{\sum_{i=1}^{N} A_{i,i}}{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{i,j}} \quad \in [0,1]\,, \tag{20}$$

where $A_{i,j}$ gives the number of pixels labelled as j and belonging to the class i in reality. Table IV shows the evolution of the classification accuracy depending on the input features. One can see also that the four shape features are not sufficient for good discrimination, whereas the four spatial features perform well. One can notice that the inclusion of the shape features improves the OA from step to step. When considering the six ROIs, the modeling of lidar waveforms allows gaining 2.3% of OA (difference between $f_{v_2} \cup \{A\}$ and $f_{v_2} \cup \{A, w, s, \mathcal{M}\}$). This is particularly due to the width parameter, that can be retrieved with a simple Gaussian assumption, but which is better estimated with the library of shapes. It leads to a better discrimination of building and vegetation areas. Indeed, for steeped roofs the four spatial feature values may be very similar to dense tree canopies. The relevance of s and $\mathcal{M}$ is lower, but the classification results benefit from their introduction.

Figure 11 gives the classification results for three ROIs. When dealing with flat homogeneous surfaces, the classifier performs well (OA=99.3%, see Table IV), however the label image looks slightly noisy. For complex mixed urban areas, the OA is satisfactory and the label images are spatially coherent (Figures 11b and c). Misclassified areas can be mainly noticed

Overall Accuracy	Area			
	6 ROIs	Grass	Building #1	Building #2
f_{v_1}	68.4	98.17	64.6	68.1
f_{v_2}	85.5	99.11	82.7	83.5
$f_{v_2} \cup \{A\}$	86.7	99.34	86.1	84.42
$f_{v_2} \cup \{A, w\}$	88.3	82.7	87.9	84.59
$f_{v_2} \cup \{A, w, s\}$	88.7	**99.35**	88.0	84.67
$f_{v_2} \cup \{A, w, s, \mathcal{M}\}$	**89.0**	99.31	**88.3**	**84.71**

TABLE IV
OVERALL ACCURACY EVOLUTION DEPENDING ON THE FEATURES INCLUDED IN THE SVM ALGORITHM.



(a) (b) (c) (d)

Fig. 11. Results of the SVM classification using the eight features (Buildings – Vegetation – Ground). **(a)** and **(b)** *Grass* and *Building #1* areas (see Figure 10 for their respective orthoimages). **(c-d)** Orthoimage ©IGN and classification of the *Building # 2* area.

at building edges and vegetated areas where both ground and off-ground objects have been mixed. Moreover, low objects on the ground such as cars may locally influence the feature values for the ground class and may lead to locally misclassified pixels (Figure 11d). These objects are then classified as *building* instead of *ground*. This result does not stem from the non spatial homogeneity in the model selection for our approach as can be seen in Figure 10.
Although the four shape features are still not sufficient, we can conclude that they allow for a better discrimination when they are fed into a SVM classifier with traditional spatial lidar features.

VI. CONCLUSION AND FUTURE WORKS

We have proposed an original method for modeling lidar waveforms by complex parametric functions. The obtained results are convincing. The fitting accuracy is better than that with conventional Gaussian waveform fitting schemes. The stochastic approach is well adapted both to locate echoes in signals and accurately describe them with parametric functions taken from an extensible and tunable model library. The algorithm has been successfully applied to waveforms from different lidar sensors, and at different spatial scales, showing its effectiveness and flexibility for various landscapes and resolutions. For medium and large-sized footprints, the chosen functions allow to adjust asymmetric peaks occurring frequently. Our approach is thus particularly relevant for such data. For small-footprints, the skewness of the echoes is less significant and shows the present limitations of our model. For only slightly asymmetric echoes, all the objects of the library are suitable and can be chosen, resulting in a sort of overfitting. There are no prior constraints on the object types in our model for neighboring echoes and echoes belonging to the same waveforms. Thus, the approach can result in non-homogeneous spatial function maps.

The potential advantages of the new approach are twofold. First, 3D points can be to accurately generated over large areas with shape descriptors that are the parameters of the modeling functions. Moreover, the 3D points can be labelled with their modeling function. By providing new features, our approach offers the possibility to improve classical lidar data classification algorithms. However, processing millions of waveforms requires a significant computing time. For our experiments, with a Macintosh Pro 8-core 2.93GHz with 6GB RAM, approximatively 50,000 waveforms can be processed in one hour.
In future works, it would be interesting to estimate automatically the weighting parameters using for instance the EM algorithm. Moreover, we should introduce, in the energy formulation, specific interactions between parametric functions of different types in order to improve local signal adjustments. Eventually, as consecutive small-footprint waveforms along a scan line and in the orthogonal directions are likely to have similar shapes, spatial interactions should also be included in the regularization term of the proposed model.

ACKNOWLEDGMENTS

LVIS data set was provided by the Laser Vegetation Imaging Sensor team in the Laser Remote Sensing Branch at NASA Goddard Space Flight Center with support from the University of Maryland, College Park.

REFERENCES

[1] J. Shan, and C. Toth (Eds.), *Topographic Laser Ranging and Scanning: Principles and Processing*, Taylor & Francis, Boca Raton, USA, 2008.
[2] C. Frueh, S. Jain, A. Zakhor, "Data Processing Algorithms for Generating Textured 3D Facade Meshes from Laser Scans and Camera Images," *International Journal of Computer Vision*, vol. 61 no. 2, pp. 159–184, 2005.
[3] J. Secord, and A. Zakhor, "Tree Detection in Urban Regions Using Aerial LiDAR and Image Data," *IEEE Geoscience and Remote Sensing Letters*, vol. 4 no. 2, pp. 196–200, 2007.
[4] M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46 no. 5, pp. 1416–1427, 2008.
[5] C. Mallet, and F. Bretar, "Full-Waveform Topographic Lidar: State-of-the-art," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 1, pp. 1–16, 2009.
[6] M.A. Hofton, J.B. Minster, and J.B. Blair, "Decomposition of Laser Altimeter Waveforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 4, pp. 1989–1996, 2000.
[7] W. Wagner, A. Ullrich, V. Ducic, T. Melzer and N. Studnicka, "Gaussian Decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 60, no. 2, pp. 100–112, 2006.
[8] M. Kirchhof, B. Jutzi, and U. Stilla, "Iterative Processing of Laser Scanning Data by Full-Waveform Analysis," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 1, pp. 99–114, 2008.
[9] C.C. Holmes, and D.G.T. Denison, "Perfect sampling for the wavelet reconstruction of signals," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 237–244, 2002.
[10] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42 no. 6, pp. 1321–1334, 2004.
[11] M. Unser, "Splines: A Perfect Fit for Signal and Image Processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, 1999.
[12] Y. Bengio, and P. Vincent, "Manifold Parzen Windows," *In Advances in Neural Information Processing Systems (NIPS)*," 15, MIT Press, pp. 825–832. 2003.

[13] J. Weston, A. Gammerman, M. Stitson, V. Vapnik, V. Vovk, and C. Watkins. "Support vector density estimation," in *Advances in Kernel Methods Support Vector Learning*, B. Scholköpf, C.J.C. Burges, and A.J. Smola (Eds), MIT Press, Cambridge, CA, USA, pp. 293–306, 1999.

[14] G.J. McLachlan, and D. Peel, "Finite Mixture Models," *Wiley Series in Probability and Statistic*, Wiley, New-York, NY, USA, 2000.

[15] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.

[16] M. Figueiredo, and A.K. Jain, "Unsupervised Learning of Finite Mixture Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381–396, 2002.

[17] E.E. Kuruoğlu, and J. Zerubia, "Modeling SAR Images With a Generalization of the Rayleigh Distribution," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 527–533, 2004.

[18] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre, "A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 10, no. 42, pp. 2046–2057, 2004.

[19] V. Krylov, G. Moser, S.B. Serpico, and J. Zerubia, "Modeling the statistics of high resolution SAR images," *INRIA Technical Report*, no 6722, INRIA Sophia-Antipolis, France, 2008.

[20] B. Jutzi and U. Stilla, "Range determination with waveform recording laser systems using a Wiener Filter," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 61, no. 2, pp. 95–107, 2006.

[21] A. Baddeley, and M. Van Lieshout, "Stochastic geometry models in high-level vision," *Statistics and Images*, vol. 1, no. 2, pp. 233–258, 1993.

[22] M.N. Van Lieshout, *Markov point processes and their applications*, Imperial College Press, London, UK, 2000.

[23] L. Holden, S. Sannan, and H. Bungum, "A stochastic marked point process model for earthquakes," *Natural Hazards and Earth System Sciences*, vol. 3, no. 1/2, pp. 95–101, 2003.

[24] C. Lacoste, X. Descombes, and J. Zerubia, "Point processes for unsupervised line network extraction in remote sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 1568–1579, 2005.

[25] O. Tournaire, and N. Paparoditis, "A geometric stochastic approach based on marked point processes for road mark detection from high resolution aerial images," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 64, no. 6, pp. 621–631, 2009.

[26] K. Sun, N. Sang, and T. Zhang, "Marked Point Process for Vascular Tree Extraction on Angiogram," in Proc. *Energy Minimization Methods in Computer Vision and Pattern Recognition*, Lecture Notes in Computer Science, vol. 4679, pp. 467–478, Springer-Verlag, Berlin- Heidelberg, Germany, 2007.

[27] G. Perrin, X. Descombes, and J. Zerubia, "Adaptive Simulated Annealing for Energy Minimization Problem in a Marked Point Process Application," in Proc. *Energy Minimization Methods in Computer Vision and Pattern Recognition*, Lecture Notes in Computer Science, vol. 3757, pp. 3–17, Springer-Verlag, Berlin- Heidelberg, Germany, 2005.

[28] M. Ortner, X. Descombes, and J. Zerubia, "Building outline extraction from Digital Elevation Models using marked point processes," *International Journal of Computer Vision*, vol. 72, no. 2, pp. 107–132, 2007.

[29] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, "Structural approach for building reconstruction from a single DSM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 135–147, 2010.

[30] W.K. Hastings, "Monte Carlo sampling using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[31] E. Punskaya, C. Andrieu, A. Doucet, and W. Fitzgerald, "Bayesian curve fitting using MCMC with applications to signal segmentation," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 747–758, 2002.

[32] S. Hernández-Marín, A. Wallace, and G. Gibson, "Bayesian Analysis of Lidar Signals with Multiple Returns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2170–2180, 2007.

[33] D. Salas-González, E.E. Kuruoğlu, and D.P. Ruiz, "Finite mixture of α-stable distributions," *Digital Signal Processing*, vol. 19, no. 2, pp. 360–369, 2009.

[34] P.J. Green, "Reversible Jump Markov Chain Monte-Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 97–109, 1995.

[35] C. Mallet, F. Lafarge, F. Bretar, U. Soergel, and C. Heipke, "Lidar Waveform Modeling using a Marked Point Process," in Proc. *IEEE International Conference on Image Processing*, pp. 1713–1716, Cairo, Egypt, 2009.

[36] H. Rue, and A. Syverseen, "Bayesian Object Recognition with Baddeleys Delta Loss," *Advances in Applied Probability*, vol. 30, no. 1, pp. 64–84, 1998.

[37] A. Pievatolo, and P.J. Green, "Boundary detection through dynamic polygons," *Journal of the Royal Statistical Society*, vol. 60, no. 3, pp. 609–626, 1998.

[38] S. Geman, and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[39] W. Ge, and R. Collins "Marked Point Processes for Crowd Counting," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, pp. 2913–2920, 2009.

[40] F. Lafarge, X. Descombes, and G. Gimel'farb, "Geometric Feature Extraction by a Multi-Marked Point Process," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

[41] N. Chehata, L. Guo and C. Mallet, "Contribution of Airborne Full-Waveform Lidar and Image Data for Urban Scene Classification," in Proc. *IEEE International Conference on Image Processing*, pp. 1669–1672, Cairo, Egypt, 2009.

[42] C. Mallet, F. Bretar, and U. Soergel, "Analysis of Full-Waveform Lidar Data for Classification of Urban Areas," *Photogrammetrie Fernerkundung Geoinformation*, vol. 5, pp. 337–349, 2008.

[43] G. Moser, J. Zerubia, and S.B. Serpico, "SAR Amplitude Probability Density Function Estimation Based on a Generalized Gaussian Model," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1429–1444, 2006.

[44] A. Dick, P. Torr, R. Cipolla, "Modelling and interpretation of architecture from several images," *International Journal of Computer Vision*, vol. 60 no. 2, pp. 111–134, 2004.

[45] N. Ripperda, "Grammar based facades reconstruction using rjMCMC," *Photogrammetrie Fernerkundung Geoinformation*, vol. 2, pp. 83–92, 2008.

[46] P. Salamon, P. Sibani, and R. Frost, "Facts, Conjectures and Improvements for Simulated Annealing," *SIAM Monographs on Mathematical Modeling and Computation*, Society for Industrial and Applied Mathematics, Philadelphia, PN, USA, 2002.

[47] M. Hofton, L. Rocchio, J. Blair, R. Dubayah, "Validation of Vegetation Canopy Lidar sub-canopy topography measurements for a dense tropical forest," *Journal of Geodynamics*, vol. 34, no. 3-4, pp. 491–502, 2002.

[48] D.J. Harding, "BOREAS Scanning Lidar Imager of Canopies by Echo Recovery (SLICER): Level-3 Data," *Available by special arrangement with Oak Ridge National Laboratory Distribued Active Archive Center*, Oak Ridge, Tennessee, U.S.A. [http://www.daac.ornl.gov], 2000 (on CD-ROM).

[49] A. Chauve, F. Bretar, S. Durrieu, M. Pierrot-Deseilligny, and W. Puech, "FullAnalyze: a Research Tool for Handling Processing and Analyzing Full-waveform Lidar Data," in Proc. *IEEE International Geoscience and Remote Sensing Symposium*, Cape Town, South Africa, 2009.

[50] B. Schölkopf, C. Burges, and A. Smola, "Advances in Kernel Methods. Support Vector Learning," *The MIT Press*, Cambridge, USA, 1998.

[51] V. Vapnik, "The Nature of Statistical Learning Theory," *Springer*, New-York, NY, USA, 1999.